

Frequency filtering method using a Wiener filter applied to noise reduction of audio signals

Publication number: EP0918317 (A1)

Publication date: 1999-05-26

Inventor(s): PASTOR DOMINIQUE [FR]; REYNAUD GERARD [FR];
BRETON PIERRE-ALBERT [FR]

Applicant(s): SEXTANT AVIONIQUE [FR]

Classification:

- international: G10L15/20; G10L21/02; H03H17/02; H04B1/10; G10L15/00;
G10L21/00; H03H17/02; H04B1/10; (IPC1-7): G10L3/02

- European: G10L21/02A1

Application number: EP19980402894 19981120

Priority number(s): FR19970014641 19971121

Also published as:

EP0918317 (B1)

US6445801 (B1)

JP11265198 (A)

FR2771542 (A1)

Cited documents:

EP0534837 (A1)

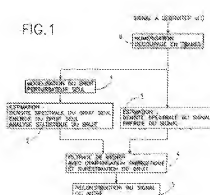
XP000658118 (A)

XP004059541 (A)

XP000658120 (A)

Abstract of EP 0918317 (A1)

The noisy signal is converted into the frequency domain by Fourier analysis (1). A model is created of the noise (2) and estimates made of the spectral density, energy level and the conversion coefficient for statistical dispersion. In parallel to this similar estimates are made of the noisy signal (3). The two results are then recombined in a Wiener filter (4) which compensates for energy levels and applies noise overestimation. The signal is then reconstructed (5).



Data supplied from the esp@cenet database — Worldwide

(19)



Europäische Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 918 317 A1

(12)

DEMANDE DE BREVET EUROPEEN

(43) Date de publication:

26.05.1999 Bulletin 1999/21

(51) Int. Cl.⁸ G10L 3/02

(21) Numéro de dépôt: 98402894.4

(22) Date de dépôt: 20.11.1998

(84) Etats contractants désignés:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE

Etats d'extension désignés:

AL LT LV MK RO SI

(72) Inventeurs:

- Pastor, Dominique, Thomson-CSF Prop. Intel.
94117 Arcueil Cedex (FR)
- Reynaud, Gérard, Thomson-CSF Prop. Intel.
94117 Arcueil Cedex (FR)
- Braton, Pierre-Albert, Thomson-CSF Prop. Intel.
94117 Arcueil Cedex (FR)

(30) Priorité: 21.11.1997 FR 9714641

(71) Demandeur: SEXTANT AVIONIQUE

78140 Vélizy Villacoublay (FR)

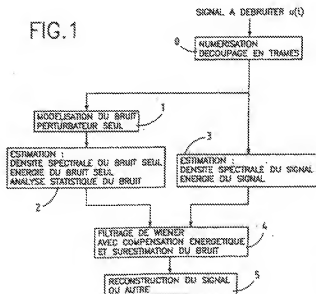
(74) Mandataire: Beylot, Jacques et al
Thomson-CSF Propriété Intellectuelle,
13, Avenue du Président Salvador Allende
94117 Arcueil Cédex (FR)

(54) Procédé de filtrage fréquentiel appliqué au débruitage de signaux sonores mettant en œuvre un filtre de Wiener

(57) Le procédé selon l'invention utilise le filtrage fréquentiel de Wiener pour débruiter des signaux sonores ($u(t)$). Il comprend une étape préliminaire consistant à numériser par échantillonnage des signaux sonores à débruiter ($u(t)$) et à les découper en trames. Le procédé comprend ensuite une première série d'étapes consistant à créer un modèle de bruit (1) sur N trames, estimer (2) la densité spectrale du bruit, l'énergie du mo-

dèle de bruit et calculer un coefficient traduisant la dispersion statistique du bruit. Il comprend également une seconde série d'étapes consistant à calculer (3) la densité spectrale des signaux à débruiter, pour chaque trame. Les coefficients du filtre de Wiener sont modifiés (4), pour chaque trame successivement traitée, à l'aide des paramètres déterminés à l'issue des deux séries d'étapes, de manière à introduire une compensation énergétique et une surestimation adaptative du bruit.

FIG. 1



Enclosure A2

Opposition re. EP 1 242 992

Qualcomm Incorporated
J.
Nokia Corporation

Description

[0001] La présente invention concerne un procédé de filtrage fréquentiel mettant en oeuvre un filtre de Wiener.

[0002] Elle s'applique notamment, bien que non exclusivement, au débruitage de signaux sonores contenant de la parole captée en milieu bruité et de façon plus générale au débruitage de tous signaux sonores.

[0003] Les domaines principaux concernent les communications téléphoniques ou radiotéléphoniques, la reconnaissance vocale, la prise de son à bord d'aéronefs civils ou militaires, et de façon plus générale de tous véhicules bruyants, les intercommunications de bord, etc.

[0004] A titre d'exemple non limitatif, dans le cas d'un aéronef, les bruits résultent des moteurs, de la climatisation, de la ventilation des équipements de bord ou des bruits aérodynamiques. Tous ces bruits sont captés, au moins partiellement, par le microphone dans lequel parle le pilote ou un autre membre de l'équipage. En outre, pour ce type d'application en particulier, une des caractéristiques des bruits est d'être très variables dans le temps. En effet, ils sont très dépendants du régime de fonctionnement des moteurs (phase de décollage, régime stabilisé, etc.). Les signaux utiles, c'est-à-dire les signaux représentant les conversations, présentent également des particularités : ils sont le plus souvent de brève durée.

[0005] Enfin, quelle que soit l'application envisagée, si on s'intéresse au "voisement", on peut mettre en évidence certaines particularités. Comme il est connu, le voisement concerne des caractéristiques élémentaires de morceaux de parole, et plus précisément concerne les voyelles, ainsi qu'une partie des consonnes : "b", "d", "g", "j", etc. Ces lettres se caractérisent par un signal audiotonique de structure pseudo-périodique.

[0006] En traitement de la parole, il est courant de considérer que les régimes stationnaires, notamment le voisement précité, s'établissent sur des durées comprises entre 10 et 20 ms. Cet intervalle de temps est caractéristique des phénomènes élémentaires de la production de la parole et sera dénommé trame ci-après.

[0007] Aussi, il est usuel que les procédés de débruitage prennent en compte cette caractéristique importante des signaux sonores comprenant de la parole.

[0008] Ces procédés comprennent généralement les étapes principales suivantes : un découpage en trames du signal audiotonique à débruiter, le traitement de ces trames par une opération de transformée de Fourier (ou d'une transformée similaire) pour passer dans le domaine fréquentiel, le traitement de débruitage proprement dit par filtrage numérique, et un traitement, dusi du premier, par une transformée de Fourier inverse, pour revenir dans le domaine temporel. La dernière étape consiste en une reconstruction du signal. Cette reconstruction peut être obtenue en multipliant chacune des trames par une fenêtre de pondération.

[0009] Un des filtres numériques les plus utilisés pour ce type d'application est le filtre de Wiener, en particulier un filtre de Wiener dit optimal. Celui-ci présente l'avantage de traiter de façon différenciée les trames successives.

[0010] En d'autres termes, et de façon plus générale, Le filtrage optimal de Wiener se trouve au centre des méthodes optimales de traitement du signal, basées sur les caractéristiques statistiques du second ordre et donc de la notion de corrélation.

[0011] Le filtrage de Wiener permet la séparation des signaux par décorrélation. Son importance est liée à la simplicité des calculs théoriques. En outre, il peut s'appliquer à une multitude de processus particuliers, et notamment, en ce qui concerne l'application préférée visée par l'invention, l'extraction d'un bruit polluant un signal de parole.

[0012] Cependant, dans l'art connu, un problème classique rencontré lors du débruitage par filtrage de Wiener est la présence d'un bruit, appelé bruit musical, qui dégrade la perception du signal débruité. Ce bruit musical est dû aux fluctuations des densités spectrales du bruit présent dans le signal d'entrée. Pour certaines trames, en effet, la densité spectrale du bruit est supérieure, au moins sur un canal fréquentiel, à celle du modèle de bruit que l'on utilise dans ces techniques. Dans ce cas, les mécanismes propres au filtrage de Wiener provoquent l'apparition d'un bruit résiduel sur le signal débruité. Celui-ci est particulièrement désagréable d'un point de vue perceptuel de part son instabilité. En effet, lors de l'écoute d'un signal de parole, on distingue des bruits résiduels sous la forme de "glouglou", qui s'apparente à des distorsions que l'on peut attribuer à une grande variabilité du bruit polluant le signal de parole débruité ou signal "utile".

[0013] L'invention se fixe donc pour but de pallier les inconvénients des procédés de filtrage de l'art connu, notamment l'inconvénient principal qui vient d'être rappelé : la présence d'un bruit résiduel parasite dans le signal débruité, dit "bruit musical". L'invention vise, de façon plus générale, à augmenter l'intelligibilité de la parole, dans son application principale.

[0014] En vue d'atténuer fortement les effets du bruit musical, l'invention tire profit des deux observations expérimentales suivantes :

- la probabilité de bruit musical est d'autant plus forte que l'estimée des densités spectrales du bruit est instable d'une trame à l'autre ;
- la probabilité de présence de bruit musical est d'autant plus forte que l'estimée de la densité spectrale du bruit est

faible par rapport à sa densité spectrale réelle.

[0015] Selon une caractéristique principale de l'invention, le filtre de Wiener utilisé pour le filtrage numérique est modifié de façon optimisée en y introduisant un terme de compensation énergétique visant à surestimer le niveau de bruit. En outre, ce terme de compensation est adaptatif.

[0016] L'invention a donc pour objet un procédé de filtrage fréquentiel pour le débruitage de signaux sonores bruités constitués de signaux sonores dits utiles mélangés à des signaux de bruit, le procédé comprenant au moins une étape de découpage desdits signaux sonores en une série de trames identiques d'une longueur déterminée et une étape de filtrage fréquentiel à l'aide d'un filtre de Wiener, caractérisé en ce qu'il comprend, en outre, les étapes suivantes :

- élaboration à partir desdits signaux bruités d'un modèle de bruit sur un nombre N déterminé desdites trames, N étant compris entre des bornes minimale et maximale prédéterminées ;
- application d'une transformée de Fourier auxdites N trames ;
- estimation, pour chaque trame dudit modèle, de la densité spectrale de cette trame ;
- estimation de la densité spectrale moyenne dudit modèle de bruit ;
- calcul, à partir de ces deux estimations, d'un coefficient de surestimation statistique, ledit coefficient statistique étant égal au rapport maximal, pour lesdites N trames du modèle de bruit, entre le maximum de la densité spectrale d'une trame considérée dudit modèle de bruit, et le maximum de la densité spectrale estimée du modèle de bruit ;
- estimation, pour chaque trame desdits signaux à débruiter, de sa densité spectrale ; et
- modification, pour chaque trame desdits signaux à débruiter, des coefficients dudit filtre de Wiener pour que la relation suivante soit vérifiée :

$$\hat{W}(v) = \left(1 - \alpha \cdot \max_i \cdot \frac{\gamma_x(v)}{\gamma_u(v)} \right)^\beta,$$

relation dans laquelle α et β sont des coefficients fixes prédéterminés, dits coefficient statique de compensation énergétique et coefficient d'atténuation exponentielle, respectivement, v décrit l'ensemble des canaux fréquentiels de ladite transformée de Fourier, $\gamma_u(v)$ étant l'estimée de la densité spectrale de la trame à débruiter, $\gamma_x(v)$ est ladite densité spectrale du modèle de bruit, et \max_i ledit coefficient de surestimation statistique, modifiant le coefficient statique de compensation énergétique α .

[0017] L'invention sera mieux comprise et d'autres caractéristiques et avantages apparaîtront à la lecture de la description qui suit en référence aux figures annexées, parmi lesquelles :

- la figure 1 illustre, sous forme de bloc diagramme, les principales étapes du procédé selon l'invention ;
- la figure 2 illustre schématiquement un filtre de Wiener de l'art connu ;
- la figure 3 est un diagramme illustrant la densité spectrale d'un modèle de bruit et les densités spectrales γ_u de chaque trame de ce modèle de bruit ;
- les figures 4a et 4b sont des diagrammes comparatifs illustrant ces mêmes paramètres avec surestimation de la densité spectrale du modèle de bruit ;
- la figure 5 est un diagramme illustrant ces mêmes paramètres avec surestimation adaptative de la densité spectrale du modèle de bruit ;
- la figure 6 représente un exemple typique de signal issu d'une prise de son bruitée ;
- la figure 7 est un organigramme représentant les étapes d'un procédé particulier de recherche d'un modèle de bruit ;

- et la figure 8 est un organigramme détaillé représentant les étapes du procédé de filtrage numérique selon un mode de réalisation préféré de l'invention.

[0018] Les principales phases et étapes du procédé selon l'invention vont maintenant être décrites par référence au bloc diagramme de la figure 1. Chaque bloc, référencés 0 à 5, représente une phase du procédé, elle-même pouvant être subdivisée en étapes élémentaires.

[0019] Dans ce qui suit, pour fixer les idées et sans que cela limite en quoi que ce soit la portée de l'invention, on va se placer, dans le cadre du traitement de la parole bruitée. Comme il a été indiqué précédemment, il est courant de considérer que les régimes stationnaires, notamment le voisement, s'établissent sur des durées comprises entre 10 et 20 ms, intervalle de temps caractéristiques des phénomènes élémentaires de la production de la parole et qui sera dénommé trame ci-après.

[0020] Comme dans l'art connu, le procédé de l'invention, comprend une étape de découpage en trames du signal audiophonique à débiter (bloc 0).

[0021] Dans la pratique, on met en oeuvre des techniques numériques. Aussi, les signaux de trame ne sont pas des signaux à "évolution continue", mais des signaux discrets, obtenus par échantillonnage. On suppose que les signaux sont échantillonnés à la période T_e , avant traitement numérique. Il est courant de considérer alors 20 échantillons pour une trame de signal, en choisissant p de manière à ce que la valeur $20 T_e$ soit de l'ordre grandeur de la durée D d'une trame. A titre d'exemple, pour une fréquence d'échantillonnage de 10 kHz, on choisit souvent des trames de 12,8 ms, de manière à pouvoir disposer de 128 points pour chaque trame, ce qui constitue une puissance de deux. Le nombre d'échantillons correspondant à une trame sera noté ci-après L (trame). La relation suivante : $D = L \text{ Trame} \times T_e$ est donc satisfaite. L'étape de découpage en trames, comme indiqué sur la figure 1, est donc précédée d'une étape de numérisation par échantillonnage.

[0022] Par convention, le signal d'entrée sera noté $u(t)$, le signal utile $s(t)$ et le bruit perturbateur $x(t)$ de telle façon que :

$$u(t) = s(t) + x(t) \text{ en temps continu} \quad (1)$$

$$u(kT_e) = s(kT_e) + x(kT_e) \text{ en temps discret} \quad (2)$$

[0023] Les étapes de numérisation et de découpage en trames (bloc 0) sont communes à l'art connu. Les échantillons numériques ainsi créés sont rangés dans une mémoire tampon circulaire de type "FIFO" (c'est-à-dire du type "premier entré - premier sorti") afin d'être lus sous forme de trames successives.

[0024] Les trames successivement lues subissent alors une série d'étapes de traitement autonomes, selon deux voies que l'on peut qualifier de "parallèles".

[0025] Les opérations effectuées dans le bloc 1, consiste à identifier des segments du signal à débiter ne contenant que du bruit. La sortie de ce bloc est constituée d'une suite d'échantillons numériques représentatifs du bruit seul. En d'autres termes, un modèle du bruit est élaboré à partir des signaux bruités, ou plus précisément à partir des trames successivement lues (bloc 0). De nombreux procédés peuvent être mis en oeuvre et un exemple du procédé de recherche de modèle de bruit sera explicité ci-après.

[0026] Dans le bloc 2, trois étapes sont réalisées et consistent, à partir des échantillons fournis par le bloc 1, à effectuer :

- l'estimation de la densité spectrale moyenne du bruit (par exemple par spectre moyen et corrélogramme liés) ;
- la détermination de l'énergie moyenne du modèle de bruit ;
- et la détermination d'un coefficient traduisant la dispersion statistique du bruit.

[0027] Les étapes ci-dessus, et notamment la dernière étape qui constitue une des caractéristiques principales de l'invention, seront détaillées ci-après.

[0028] Dans la branche "parallèle", le bloc 3 comporte une étape d'estimation de la densité spectrale de la trame courante de signal et de calcul de son énergie.

[0029] Dans le bloc 4, selon une autre caractéristique essentielle de l'invention, les coefficients du filtre fréquentiel effectuant le débruitage du signal sont déterminés de la manière qui sera détaillée ci-après. Comme il a été indiqué, le procédé de l'invention est basé sur une compensation énergétique et une surestimation du bruit.

[0030] Enfin, dans le bloc 5, le signal temporel débruité est reconstruit, en assurant la meilleure continuité possible entre les trames. Dans d'autres applications que l'application principale visée par l'invention les signaux peuvent être exploités tels quels par des divers procédés tels que la reconnaissance automatique de la parole. En soi, cette phase du procédé est commune à l'art connu, et il n'y a pas lieu de détailler la méthode de reconstruction ou d'exploitation des signaux en sortie du bloc 4.

[0031] Selon la caractéristique principale de l'invention, le procédé permet de modifier et d'optimiser les coefficients du filtre de Wiener utilisé pour la phase de débruitage proprement dite (bloc 4), de façon à éliminer ou, pour le moins, fortement atténuer, les bruits parasites dits "musicaux".

[0032] Comme il a été rappelé, ces bruits sont attribuables à deux causes principales :

a/ la probabilité de bruit musical est d'autant plus forte que l'estimée des densités spectrales du bruit est instable d'une trame à l'autre ;

b/ la probabilité de présence de bruit musical est d'autant plus forte que l'estimée de la densité spectrale du bruit est faible par rapport à la densité spectrale réelle du bruit.

[0033] Selon l'invention, en relation avec la cause a/, la dispersion est quantifiée par un coefficient issu de l'analyse effectuée dans le bloc 2, à partir du modèle de bruit élaboré dans le bloc 1.

[0034] De même, en relation avec la cause b/, pour réduire l'influence de la densité spectrale du bruit, en particulier lorsqu'elle est instable, le procédé selon l'invention effectue une surestimation de cette densité spectrale, en y introduisant un degré d'adaptivité afin d'optimiser la perception du signal débruité.

[0035] Avant de décrire plus en détail le procédé de l'invention, il est utile de rappeler brièvement les caractéristiques d'un filtre de Wiener selon l'art connu.

[0036] La figure 2 illustre de façon très schématisée un filtre de Wiener utilisé pour débruiter un signal bruité $U(n)$.

[0037] A titre d'exemples non limitatifs, des filtres de Wiener sont décrits dans les livres suivants, auxquels on pourra se référer avec profit :

- Yves THOMAS : "Signaux et systèmes linéaires", éditions MASSON (1994) ; et :

- François MICHAUD : "Méthodes adaptatives pour le signal", édition HERMES (1992).

[0038] Sur la figure 2 les conventions suivantes ont été adoptées :

- $U(n)$: transformée de Fourier discrète du processus aléatoire observé, soit le signal bruité ;
- $S(n)$: transformée de Fourier discrète du processus "désiré", à estimer par filtrage linéaire de $U(n)$;
- $X(n)$: transformée de Fourier discrète du bruit additif polluant le signal utile ;
- $\hat{S}(n)$: estimation de $S(n)$ exprimée dans le domaine de

Fourier, avec $e = \hat{S} - S$ = erreur d'estimation (S étant le signal débruité réel) ; et

- $W(z)$: filtre d'estimation exprimé dans le domaine fréquentiel.

[0039] Le filtre optimal de Wiener minimise la distance entre les variables aléatoires $S(n)$ et $\hat{S}(n)$ mesurée par l'erreur quadratique moyenne J :

$$J = E\{[S(n) - \hat{S}(n)]^2\} \quad (3)$$

[0040] La minimisation de ce critère revient à rendre l'erreur d'estimation orthogonale au signal observé, ce qui se traduit par le principe d'orthogonalité :

$$E\{e(n) \cdot U^*(n)\} = 0 \quad (4)$$

[0041] En notant :

γ_S la densité spectrale du signal utile, et
 γ_X la densité spectrale du bruit parasite.

le filtre de Wiener est décrit par la relation suivante :

$$W(n) = \frac{\gamma_S(n)}{\gamma_S(n) + \gamma_X(n)} \quad (5)$$

[0042] En prenant en compte l'indépendance de $S(n)$ et de $X(n)$, on obtient la relation ci-dessous :

$$\gamma_U = \gamma_S + \gamma_X \quad (6)$$

relation dans laquelle γ_U représentant la densité spectrale du signal observé.

[0043] La relation décrivant le filtre de Wiener devient donc finalement :

$$W(n) = \frac{\gamma_S(n)}{\gamma_S(n) + \gamma_X(n)} = 1 - \frac{\gamma_X(n)}{\gamma_U(n)} \quad (7)$$

[0044] En pratique, c'est cette seconde formulation du filtre de Wiener qui est utilisée, puisqu'elle ne fait intervenir que des termes directement accessibles, c'est-à-dire, d'une part, le signal bruité reçu du bloc 3 et, d'autre part, le bruit, préalablement déterminé par le calcul du modèle de bruit (bloc 1).

[0045] Il doit être remarqué que les coefficients $W(n)$ du filtre de Wiener sont toujours positifs. Si des artefacts de calcul provoquent une valeur négative pour un coefficient, ce coefficient est rendu égal à zéro.

[0046] Selon l'art conçu, la suppression du bruit additif par une méthode de soustraction spectrale, telle qu'elle est réalisée par un filtre Wiener, débouche sur la création de bruits dits "musicaux". Pour éviter l'apparition de ces bruits parasites désagréables à l'écoute et nuisibles à l'intelligibilité de la parole, ou pour le moins empêcher au maximum leur apparition, selon une caractéristique essentielle de l'invention, les coefficients du filtre de Wiener sont modifiés à l'aide de paramètres déterminés dans les blocs 2 et 3, de la manière qui va maintenant être détaillée.

[0047] Lorsque le signal d'entrée ne contient que du bruit, le "bruit musical" supplémentaire est présent parce que, dans la pratique, l'estimation du rapport γ_S fluctue à chaque fréquence, bien qu'en théorie ce rapport devrait être égal à l'unité toutes les fois que les fréquences γ_S sont nulles. Ce sont ces erreurs d'estimation qui produisent des filtres atténuateurs dont les variations des coefficients sont aléatoires, selon les fréquences et au cours du temps.

[0048] Pour fixer les idées, on considère l'exemple du débruitage d'un bruit seul, échantillonné à 44 kHz. On détermine la densité spectrale γ_U d'un modèle de bruit choisi à l'aide de ce signal et les densités spectrales γ_U de chaque trame (de longueur L_{Gtrame}) de ce bruit.

[0049] On a représenté la variation de ces deux paramètres sous forme de courbes dans le diagramme de la figure 3, en fonction du nombre de canaux de transformée de Fourier FFT. Pour tracer les courbes, il a été supposé que la longueur de trame était de 128 échantillons, soit $L_{Gtrame}=128$.

[0050] Ce diagramme montre clairement que les allures des deux courbes γ_U et γ_X sont similaires mais les deux estimées présentent une différence d'amplitude nette. Le pic principal de γ_U qui se situe à la fréquence 2,75 kHz (84 canaux FFT correspondant à 22 kHz, soit la demi-fréquence d'échantillonnage) a une amplitude environ sept fois supérieure à celui de γ_X situé à la même fréquence. Ceci constitue la raison principale de la présence des bruits "musicaux". Lorsque, pour certaines fréquences référencées $\gamma_U(v)$ est bien supérieur à $\gamma_X(v)$, cela signifie, en théorie, que la trame ne contient pas seulement du bruit mais une autre partie de signal. Dans ce cas, le filtrage de Wiener selon l'art connu débruite la trame correspondante comme si elle contenait du signal de parole utile, ce qui entraîne la présence de résidus de bruits.

[0051] Pour éviter cet effet parasite, le procédé selon l'invention modifie de façon optimisée les coefficients du filtre de Wiener et introduit un terme de compensation énergétique, venant surestimer artificiellement le niveau du bruit, avec différents niveaux d'adaptativité de cette compensation.

[0052] Les coefficients du filtre de Wiener modifié obéissent à la relation suivante :

$$W(v) = \left(1 - \alpha \cdot \max_i \left(\frac{E_x}{E_u} \cdot \frac{\gamma_u(v)}{\gamma_u(v)} \right) \right)^\beta \quad (8)$$

5 [0053] En se reportant de nouveau à la relation (7), on constate aisément que quatre nouveaux termes ont été introduits, à savoir :

β : coefficient d'atténuation exponentielle ;

10 α , coefficient statique de compensation énergétique ;

$\frac{E_x}{E_u}$, rapport de pondération énergétique ; et

15 \max_i : coefficient de surestimation statistique issu de l'analyse statistique du bruit, ce à partir d'un modèle de bruit établi lors de la phase du procédé correspondant au bloc 1.

[0054] Chacun de ces termes va maintenant être explicité.

20 [0055] Le coefficient d'atténuation exponentielle β est un terme communément utilisé dans la littérature consacré au domaine du filtrage numérique et, plus particulièrement, au débruitage. Une valeur typique de ce paramètre est de 0,5.

[0056] A titre d'exemple non limitatif, on pourra se reporter à l'article de L. Arsan, A. Mc Crea et V. Wiswana - Than, intitulé "New Methods for adaptive noise suppression", IEEE, mai 1995, pages B12-B15.

25 [0057] Le coefficient de compensation énergétique statique α permet de surestimer le bruit et est particulièrement pertinent dans le cas de la suppression de bruit seul. En effet, une valeur typique de $\alpha=10$ appliquée à l'exemple de la figure 3 augmente l'estimée du spectre moyen de bruit γ_u d'environ +10 dB, ce qui permet alors de diminuer le niveau de bruit résiduel, puisque les coefficients du filtre de Wiener ne peuvent être négatifs. Dans le cas contraire, ils sont alors forcés à zéro.

30 [0058] Cependant, si cette modification est très efficace pour éliminer le bruit seul, elle pose à son tour des problèmes lorsque les trames à débruirer contiennent du signal utile. Si ce signal utile est beaucoup plus énergétique que le bruit, ce coefficient multiplicateur α n'a pas d'effet sur la dégradation de ce signal. Mais, dans le cas contraire, il peut exister des fréquences v pour lesquelles une trame de signal utile a une énergie non négligeable mais proche de celle du bruit pour les mêmes fréquences. Dans ce cas, la multiplication par α de $\gamma_u(v)$ impose des coefficients de Wiener $W(v)$ nuis et donc entraîne une disparition de l'énergie du signal pour ces fréquences.

35 [0059] Ce problème est illustré par les figures 4a et 4b. Sur ces figures les conventions suivantes ont été adoptées.

γ_u : densité spectrale de la trame de signal considérée (trame de signal faiblement énergétique devant le bruit) ; et
 γ_x : densité spectrale du modèle de bruit choisi (bloc 1).

40 [0060] La courbe de la figure 4a permet de constater que l'énergie du signal dans la bande de fréquences Δv , représentée par la densité spectrale γ_u , n'est pas négligeable.

[0061] Et se référant à la figure 4b, on peut constater que la multiplication de γ_u par le paramètre $\alpha=10$ rend $\alpha \cdot \gamma_u$ supérieur à γ_u dans la bande Δv . Il s'ensuit que le gain de Wiener est nul pour cette bande de fréquences qui n'appartient plus dans la trame débruitée.

45 [0062] Le rapport de pondération énergétique décrit ci-dessous permet de réduire cette distorsion dans le signal débruité.

[0063] Comme indiqué précédemment, le débruitage du bruit seul est correct, mais il peut être trop brutal dans les parties du signal utile.

50 [0064] Dans une variante préférée de l'invention, on remédie à cet inconvénient en faisant varier le coefficient α , ce en fonction de la présence ou non d'une partie de signal utile dans le signal à débruirer. De façon avantageuse, α reste proche d'une valeur typique égale à 10, lorsque le signal bruité ne contient que du bruit, et varie entre 0 et 10, lorsqu'un signal utile est présent dans le signal bruité. On introduit donc avantageusement un degré d'adaptativité.

[0065] C'est la fonction qui est assignée au rapport $\frac{E_x}{E_u}$ qui vient multiplier α dans la relation (8), rapport dans lequel E_x est l'énergie moyenne du modèle de bruit et E_u l'énergie de la trame courante. Cela permet donc aux coefficients du filtre de Wiener de changer à chaque trame de façon différenciée selon la présence plus ou moins grande (en terme d'énergie) du signal de parole.

[0066] Si $E_x = E_u$, alors $\alpha=10$ et la trame est considérée comme du bruit seul. Elle est correctement débruitée.

55 [0067] Si au contraire $E_x \ll E_u$, cela signifie que la trame considérée est très énergétique devant le bruit et qu'il est nécessaire d'atténuer au minimum cette partie de signal.

[0068] Cette troisième modification est illustrée par la figure 5. Sur cette figure, la trame de signal considérée est la même que celle utilisée pour les figures 4a et 4b,

$$\alpha = 10 \text{ et } \frac{\sum_{i=1}^N \gamma_i}{\sum_{i=1}^N \gamma_x} = 0,2.$$

[0069] Grâce à cette pondération du coefficient α par $\frac{\sum_{i=1}^N \gamma_i}{\sum_{i=1}^N \gamma_x}$, la bande de fréquences $\Delta\nu$ dans laquelle le signal utile est éliminé, (c'est-à-dire les fréquences pour lesquelles les coefficients de γ_i sont supérieurs à ceux de γ_x) est bien moins importante que lors de la modification par multiplication du seul coefficient $\alpha=10$.

[0070] Ce type de filtre présente donc une bonne efficacité en termes d'élimination des segments du signal dégradés dans lesquels la parole est absente et de diminution des distorsions infligées au signal de parole utile.

[0071] La probabilité de génération du "bruit musical" est également liée, comme il a été indiqué, à la variance des estimées de la densité spectrale du bruit sur l'ensemble des trames.

[0072] En effet, plus les densités spectrales estimées du bruit varient d'une trame à l'autre, plus la formation du bruit "musical" est probable.

[0073] Selon un autre aspect important de l'invention, on rend dépendant la valeur du coefficient de surestimation des propriétés statistiques du bruit. Pour ce faire, il est introduit un coefficient, appelé *maxi* ci-après, proportionnel à la dispersion des valeurs de densités spectrales du bruit.

[0074] Le coefficient de surestimation devient alors :
 $\alpha = \alpha \cdot \text{maxi}$, avec *maxi* satisfaisant la relation suivante :

$$\text{maxi} = \text{Max}_{i=1..N} \left(\frac{\text{Max}(\gamma_i(\nu))}{\text{Max}(\gamma_x(\nu))} \right) \quad (9),$$

relation dans laquelle :

- *N* est le nombre de trames du modèle de bruit ;
- ν décrit l'ensemble des canaux fréquentiels, soit $L \cdot \text{Gtrame}/2$ canaux ;
- $\gamma_i(\nu)$ est la densité spectrale de la *i*ème trame du modèle de bruit dans le canal ν ; et
- $\gamma_x(\nu)$ est la densité spectrale du modèle de bruit.

[0075] Le coefficient *maxi* est égal au rapport maximal, pour toutes les trames du modèle de bruit, entre le maximum de la densité spectrale de la trame du modèle de bruit considérée, et le maximum de la densité spectrale estimée du modèle de bruit.

[0076] En d'autres termes, ce coefficient caractérise la disparité maximale du bruit pour les canaux fréquentiels portant une énergie importante. Multiplié par le coefficient α , il apporte une atténuation complémentaire proportionnelle à cette disparité.

[0077] Pour élaborer une partie des paramètres entrant dans la modification des coefficients du filtre de Wiener, il est nécessaire de disposer d'un modèle de bruit (bloc 1 de la figure 1).

[0078] L'élaboration d'un modèle de bruit d'un signal bruité est une opération classique en soi. Cependant, la méthode spécifique mise en oeuvre pour cette opération peut être une méthode de l'art connu, mais aussi une méthode originale.

[0079] On va décrire ci-après, par référence aux figures 6 et 7, une méthode d'élaboration d'un modèle de bruit, particulièrement adaptée aux applications principales visées par le procédé de l'invention, notamment le débruitage de signaux de parole bruités.

[0080] La méthode repose sur une recherche permanente et automatique d'un modèle de bruit. Cette recherche est faite sur les échantillons de signal $u(t)$ numérisés et stockés dans une mémoire tampon d'entrée. Cette mémoire est capable de mémoriser simultanément tous les échantillons de plusieurs trames du signal d'entrée (au moins 2 trames et, dans le cas général, *N* trames).

[0081] Le modèle de bruit recherché est constitué par une succession de plusieurs trames dont la stabilité en énergie et le niveau d'énergie relative font penser qu'il s'agit d'un bruit ambiant et non d'un signal de parole ou d'un autre bruit

parturbateur. On verra plus loin comment se fait cette recherche automatique.

[0082] Lorsqu'un modèle de bruit est trouvé, tous les échantillons des N trames successives représentant ce modèle de bruit sont conservés en mémoire, de sorte que le spectre de ce bruit peut être analysé et peut servir au débruitage. Mais la recherche automatique de bruit continue à partir du signal d'entrée $u(t)$ pour trouver éventuellement un modèle plus récent et plus adapté, soit parce qu'il représente mieux le bruit ambiant, soit parce que le bruit ambiant a évolué. Le modèle de bruit plus récent est mis en mémoire à la place du précédent, si la comparaison avec le précédent montre qu'il est plus représentatif du bruit ambiant.

[0083] Les postulats de départ pour l'élaboration automatique d'un modèle de bruit sont les suivants :

- le bruit qu'on veut éliminer est le bruit de fond ambiant,
- le bruit ambiant a une énergie relativement stable à court terme,
- la parole est le plus souvent précédée d'un bruit de respiration du pilote qu'il ne faut pas confondre avec le bruit ambiant; mais ce bruit de respiration s'éteint quelques centaines de millisecondes avant la première émission de parole proprement dite, de sorte qu'on ne retrouve que le bruit ambiant juste avant l'émission de parole,
- et enfin, les bruits et la parole se superposent en termes d'énergie de signal, de sorte qu'un signal contenant de la parole ou un bruit perturbateur, y compris la respiration dans le microphone, contient forcément plus d'énergie qu'un signal de bruit ambiant.

[0084] Il en résulte qu'on fera l'hypothèse simple suivante : le bruit ambiant est un signal présentant une énergie minimale stable à court terme. Par court terme, il faut entendre quelques trames, et on verra dans l'exemple pratique donné ci-après que le nombre de trames destiné à évaluer la stabilité du bruit est de 5 à 20. L'énergie doit être stable sur plusieurs trames, faute de quoi on doit supposer que le signal contient plutôt de la parole ou un bruit autre que le bruit ambiant. Elle doit être minimale, faute de quoi on considère que le signal contient de la respiration ou des éléments phonétiques de parole ressemblant à du bruit mais se superposant au bruit ambiant.

[0085] La figure 6 représente une configuration typique d'évolution temporelle de l'énergie d'un signal microphonique au moment d'un début d'émission, de parole, avec une phase de bruit de respiration, qui s'éteint pendant quelques dizaines à centaines de millisecondes pour faire place au bruit ambiant seul, après quoi un niveau d'énergie élevé indique la présence de parole, pour revenir enfin au bruit ambiant.

[0086] La recherche automatique du bruit ambiant consiste alors à trouver au moins N trames successives (par exemple $N = 5$) dont les énergies sont proches les unes des autres, c'est-à-dire que le rapport entre l'énergie de signal contenue dans une trame et l'énergie de signal contenue dans la ou, de préférence, les trames précédentes est situé à l'intérieur d'une gamme de valeurs déterminées (par exemple compris entre 1/3 et 3). Lorsqu'une telle succession de trames d'énergie relativement stable a été trouvée, on stocke les valeurs numériques de tous les échantillons de ces N trames. Cet ensemble de $N \times P$ échantillons constitue le modèle courant du bruit. Il est utilisé dans le débruitage. L'analyse des trames suivantes continue. Si on trouve une autre succession d'au moins N trames successives répondant aux mêmes conditions de stabilité d'énergie (rapports d'énergies de trames dans une gamme déterminée), on compare alors l'énergie moyenne de cette nouvelle succession de trames à l'énergie moyenne du modèle stocké, et on remplace ce dernier par la nouvelle succession si le rapport entre l'énergie moyenne de la nouvelle succession et l'énergie moyenne du modèle stocké est inférieur à un seuil de remplacement déterminé qui peut être de 1,5 par exemple.

[0087] Or ce remplacement d'un modèle de bruit par un modèle plus récent moins énergétique ou pas beaucoup plus énergétique, il résulte que le modèle de bruit se cale globalement sur le bruit ambiant permanent. Même avant une prise de parole, précédée d'une respiration, il existe une phase où le bruit ambiant seul est présent pendant une durée suffisante pour pouvoir être pris en compte comme modèle de bruit actif. Cette phase de bruit ambiant seul, après respiration, est brève. Le nombre N est choisi relativement faible, afin qu'on ait le temps de recaler le modèle de bruit sur le bruit ambiant après la phase de respiration.

[0088] Si le bruit ambiant évolue lentement, l'évolution sera prise en compte du fait que le seuil de comparaison avec le modèle stocké est supérieur à 1. S'il évolue plus rapidement dans le sens croissant, l'évolution risque de ne pas être prise en compte, de sorte qu'il est préférable de prévoir de temps en temps une réinitialisation de la recherche d'un modèle de bruit. Par exemple, dans un avion au sol à l'arrêt, le bruit ambiant sera relativement faible, et il ne faudrait pas qu'au cours de la phase de décollage le modèle de bruit reste figé sur ce qu'il était à l'arrêt du fait qu'un modèle de bruit n'est remplacé que par un modèle moins énergétique ou pas beaucoup plus énergétique. On expliquera plus loin les méthodes de réinitialisation envisagées.

[0089] La figure 7 représente un organigramme des opérations de recherche automatique d'un modèle de bruit ambiant.

[0090] Le signal d'entrée $u(t)$, échantillonné à la fréquence $F_s = 1/T_s$ et numérisé par un convertisseur analogique-numérique, est stocké dans une mémoire tampon capable de stocker tous les échantillons d'au moins 2 trames.

[0091] Le numéro de la trame courante dans une opération de recherche d'un modèle de bruit est désigné par n et

est complé par un compteur au fur et à mesure de la recherche. A l'initialisation de la recherche, n est mis à 1. Ce numéro n sera incrémenté au fur et à mesure de l'élaboration d'un modèle de plusieurs trames successives. Lorsqu'on analyse la trame courante n , le modèle comprend déjà par hypothèse $n-1$ trames successives répondant aux conditions imposées pour faire partie d'un modèle.

[0092] On considère d'abord qu'il s'agit d'une première élaboration de modèle, aucun autre modèle précédent n'ayant été construit. On verra ensuite ce qui se passe pour des élaborations ultérieures.

[0093] L'énergie de signal de la trame est calculée par sommation des carrés des valeurs numériques des échantillons de la trame. Elle est conservée en mémoire.

[0094] On lit ensuite la trame suivante de rang $n = 2$, et son énergie est calculée de la même manière. Elle est également conservée en mémoire.

[0095] On calcule le rapport entre les énergies des deux trames. Si ce rapport est compris entre deux seuils S et S' dont l'un est supérieur à 1 et l'autre est inférieur à 1, on considère que les énergies des deux trames sont proches et que les deux trames peuvent faire partie d'un modèle de bruit. Les seuils S et S' sont de préférence inverses l'un de l'autre ($S = 1/S'$) de sorte qu'il suffit de définir l'un pour avoir l'autre. Par exemple, une valeur typique est $S = 3$, $S' = 1/3$. Si les trames peuvent faire partie d'un même modèle de bruit, les échantillons qui les composent sont stockés pour commencer à construire le modèle, et la recherche continue par itération en incrémentant n d'une unité.

[0096] Si le rapport entre les énergies des deux premières trames sort de l'intervalle imposé, les trames sont déclarées incompatibles et la recherche est réinitialisée en remettant n à 1.

[0097] Dans le cas où la recherche continue, on incrémente le rang n de la trame courante, et on effectue, dans une boucle de procédure itérative, un calcul d'énergie de la trame suivante et une comparaison avec l'énergie de la trame précédente ou des trames précédentes, en utilisant les seuils S et S' .

[0098] On note à ce propos que deux types de comparaison sont possibles pour ajouter une trame à $n-1$ trames précédentes qui ont déjà été considérées comme homogènes en énergie : le premier type de comparaison consiste à comparer uniquement l'énergie de la trame n à l'énergie de la trame $n-1$. Le deuxième type consiste à comparer l'énergie de la trame n à chacune des trames 1 à $n-1$. La deuxième manière aboutit à une plus grande homogénéité du modèle mais elle a l'inconvénient de ne pas prendre en compte suffisamment bien les cas où le niveau de bruit croît ou décroît rapidement.

[0099] Ainsi, l'énergie de la trame de rang n est comparée avec l'énergie de la trame de rang $n-1$ et éventuellement d'autres trames précédentes (pas forcément toutes d'ailleurs).

[0100] Si la comparaison indique qu'il n'y a pas homogénéité avec les trames précédentes, du fait que le rapport des énergies n'est pas compris entre $1/S$ et S , deux cas sont possibles :

- ou bien n est inférieur ou égal à un nombre minimal $N1$ en dessous duquel le modèle ne peut pas être considéré comme significatif du bruit ambiant parce que la durée d'homogénéité est trop courte, par exemple $N1 = 5$; dans ce cas on abandonne le modèle en cours d'élaboration, et on réinitialise la recherche au début en remettant n à 1;

- ou bien n est supérieur au nombre minimal $N1$. Dans ce cas, puisqu'on trouve maintenant un manque d'homogénéité, on considère qu'il y a peut-être un début de parole après une phase de bruit homogène, et on conserve à titre de modèle de bruit tous les échantillons des $n-1$ trames de bruit homogènes qui ont précédé le manque d'homogénéité. Ce modèle reste stocké jusqu'à ce qu'on trouve un modèle plus récent qui semble également représenter du bruit ambiant. La recherche est réinitialisée de toute façon en remettant n à 1.

[0101] Mais la comparaison de la trame n avec les précédentes aurait pu encore aboutir à la constatation d'une trame encore homogène en énergie avec la ou les précédentes. Dans ce cas, ou bien n est inférieur à un deuxième nombre $N2$ (par exemple $N2 = 30$) qui représente la longueur maximale souhaitée pour le modèle de bruit, ou bien n est devenu égal à ce nombre $N2$. Le nombre $N2$ est choisi de manière à limiter le temps de calcul dans les opérations ultérieures d'estimation de densité spectrale de bruit.

[0102] Si n est inférieur à $N2$, la trame homogène est ajoutée aux précédentes pour contribuer à construire le modèle de bruit, n est incrémenté et la trame suivante est analysée.

[0103] Si n est égal à $N2$, la trame est également ajoutée aux $n-1$ trames homogènes précédentes et le modèle de n trames homogènes est stocké pour servir dans l'élimination du bruit. La recherche d'un modèle est par ailleurs réinitialisée en remettant n à 1.

[0104] Les étapes précédentes concernent la première recherche de modèle. Mais une fois qu'un modèle a été stocké, il peut à tout moment être remplacé par un modèle plus récent.

[0105] La condition de remplacement est encore une condition d'énergie, mais cette fois elle porte sur l'énergie moyenne du modèle et non plus sur l'énergie de chaque trame.

[0106] Par conséquent, si un modèle possible vient d'être trouvé, avec N trames où $N1 < N < N2$, on calcule l'énergie moyenne de ce modèle qui est la somme des énergies des N trames, divisée par N , et on la compare à l'énergie

moyenne des N trames du modèle précédemment stocké.

[0107] Si le rapport entre l'énergie moyenne du nouveau modèle possible et l'énergie moyenne du modèle actuel en vigueur est inférieur à un seuil de remplacement SR , le nouveau modèle est considéré comme meilleur et on le stocke à la place du précédent. Sinon, le nouveau modèle est rejeté et l'ancien reste en vigueur.

[0108] Le seuil SR est de préférence légèrement supérieur à 1.

[0109] Si le seuil SR était inférieur ou égal à 1, on stockerait à chaque fois les trames homogènes les moins énergétiques, ce qui correspond bien au fait qu'on considère que le bruit ambiant est le niveau d'énergie au dessous duquel on ne descend jamais. Mais, on éliminerait toute possibilité d'évolution du modèle si le bruit ambiant se mettait à augmenter.

[0110] Si le seuil SR était trop élevé au dessus de 1, on risquerait de mal distinguer le bruit ambiant et d'autres bruits perturbateurs (respiration), voire même certains phonèmes qui ressemblent à du bruit (consonnes sifflantes ou chuintantes par exemple). L'élimination de bruit à partir d'un modèle de bruit calé sur la respiration ou sur des consonnes sifflantes ou chuintantes requerrait alors de nuire à l'intelligibilité du signal débruité.

[0111] Dans un exemple préféré le seuil SR est d'environ 1,5. Au-dessus de ce seuil on conservera l'ancien modèle ; en dessous de ce seuil on remplacera l'ancien modèle par le nouveau. Dans les deux cas, on réinitialisera la recherche en recommençant la lecture d'une première trame du signal d'entrée $u(t)$, et en mettant n à 1.

[0112] Pour rendre l'élaboration du modèle de bruit plus fiable, on peut prévoir que la recherche d'un modèle est inhibée si une émission de parole est détectée dans le signal utile. Les traitements numériques de signal couramment utilisés en détection de parole permettent d'identifier la présence de paroles en se fondant sur les spectres caractéristiques de périodicité de certains phonèmes, notamment les phonèmes correspondant à des voyelles ou à des consonnes voisées.

[0113] Le but de cette inhibition est d'éviter que certains sons soient pris pour du bruit, alors que ce sont des phonèmes utiles, qu'un modèle de bruit fondé sur ces sons soit stocké et que la suppression du bruit postérieure à l'élaboration du modèle tende alors à supprimer tous les sons similaires.

[0114] Par ailleurs, il est souhaitable de prévoir de temps en temps une réinitialisation de la recherche du modèle pour permettre une remise à jour du modèle alors que les augmentations du bruit ambiant n'ont pas été prises en compte du fait que SR n'est pas beaucoup supérieur à 1.

[0115] Le bruit ambiant peut en effet augmenter de façon importante et rapide, par exemple pendant la phase d'accélération des moteurs d'un avion ou d'un autre véhicule, aérien, terrestre ou maritime. Mais le seuil SR impose que le modèle de bruit précédent soit conservé lorsque l'énergie moyenne de bruit augmente trop vite.

[0116] Si on souhaite remédier à cette situation, on peut procéder de différentes manières, mais la manière la plus simple est de réinitialiser le modèle périodiquement en recherchant un nouveau modèle et en l'imposant comme modèle actif indépendamment de la comparaison entre ce modèle et le modèle précédemment stocké. La périodicité peut être basée sur la durée moyenne d'élocution dans l'application envisagée ; par exemple les durées d'élocution sont en moyenne de quelques secondes pour l'équipage d'un avion, et la réinitialisation peut avoir lieu avec une périodicité de quelques secondes.

[0117] La mise en œuvre de la méthode d'élaboration d'un modèle de bruit (figure 1 : bloc 1), et, de façon plus générale du procédé selon l'invention, peut se faire à partir de calculateurs non spécialisés, pourvus de programmes de calcul nécessaires et recevant les échantillons de signaux numérisés tels qu'ils sont fournis par un convertisseur analogique-numérique, via un port adapté.

[0118] Cette mise en œuvre peut aussi se faire à partir d'un calculateur spécialisé à base de processeurs de signaux numériques, ce qui permet de traiter plus rapidement un plus grand nombre de signaux numériques.

[0119] Les calculateurs sont associés, comme il est bien connu, à différents types de mémoires, statiques et dynamiques, pour enregistrer les programmes et les données intermédiaires, ainsi qu'à des mémoires circulantes de type "FIFO". Le système comprend enfin un convertisseur analogique-numérique, pour la numérisation des signaux $u(t)$, et un convertisseur numérique-analogique, en tant que de besoin, si les signaux débruités doivent être utilisés sous forme analogiques.

[0120] En conclusion, et pour décrire de façon plus détaillée le procédé de l'invention, on peut découper les étapes de façon différente de ce qui a été décrit en référence à la figure 1 (qui illustre le procédé de façon plus synthétique). La figure 9 est un diagramme résumant toutes les étapes du procédé de filtrage selon l'invention, dans un mode de réalisation préféré.

[0121] Ces étapes se répartissent en un premier sous-ensemble d'étapes permettant de déterminer les paramètres dépendant du modèle de bruit, et un second sous-ensemble d'étapes permettant de déterminer les paramètres dépendant seulement de la trame courante du signal à débruitier.

[0122] La première étape du premier sous-ensemble, comprend une étape initiale de sélection d'un modèle de bruit adapté à l'application spécifique, avantageusement un modèle de bruit déterminé par la méthode décrite ci-dessus, en référence aux figures 6 et 7.

[0123] Ce premier sous-ensemble d'étapes comprend deux branches.

[0124] Dans la première branche, on calcule pour chaque trame du modèle de bruit (dans le domaine temporel), l'énergie de la trame, puis on calcule l'énergie moyenne des trames du modèle, ce qui permet d'estimer l'énergie moyenne du modèle, c'est-à-dire le paramètre E_w .

[0125] Dans la deuxième branche, on applique une transformée de Fourier aux trames du modèle de bruit, de façon à passer dans le domaine fréquentiel. Puis on détermine successivement la densité spectrale de la trame i (avec $i = 1 \dots N$) du modèle de bruit dans le canal fréquentiel v , soit $\gamma_i(v)$, et la densité spectrale du modèle de bruit dans le canal fréquentiel v , soit $\gamma_w(v)$. A partir de ces deux paramètres, on détermine le coefficient statistique \max_i de telle sorte qu'il vérifie la relation (9). Le paramètre $\gamma_w(v)$ est également utilisé pour le calcul d'un des autres coefficients du filtre de Wiener.

[0126] Le second sous-ensemble d'étapes comprend également deux branches.

[0127] Dans la première branche, on détermine l'énergie de la trame courante, soit E_u , et dans la seconde branche, on effectue l'estimation de la densité spectrale de la trame courante γ_u .

[0128] A partir de ces deux paramètres et des paramètres γ_w et E_w déterminés précédemment, on obtient les coefficients $\{E_u/E_w\}$ et $\{\gamma_u(v)/\gamma_w(v)\}$.

[0129] Tous les coefficients du filtre de Wiener, conforme à la relation (8), sont donc déterminés à l'issue de ces étapes. Les coefficients α et β sont des coefficients fixes prédéterminés, typiquement égaux à 10 et 0,5, respectivement.

[0130] A la lecture de ce qui précède, on constate aisément que l'invention atteint bien les buts qu'elle s'est fixés.

[0131] Il doit être clair cependant que l'invention n'est pas limitée aux seuls exemples de réalisations explicitement décrits, notamment en relation avec les figures 1 à 8.

[0132] En particulier, les exemples numériques n'ont été donnés que pour mieux préciser l'invention mais sont essentiellement liés à l'application spécifique envisagée. De ce fait, ils comportent d'un simple choix technologique à la portée de l'Homme du Métier.

[0133] En outre, comme il été rappelé, l'invention ne se réduit pas au seul domaine du filtrage de signaux contenant de la parole bruitée, même si ce domaine constitue une des applications préférées.

Revendications

1. Procédé de filtrage fréquentiel pour le débruitage de signaux sonores bruités ($\{u(i)\}$) constitués de signaux sonores dits utiles mélangés à des signaux de bruit, le procédé comprenant au moins une étape de découpage (0) desdits signaux sonores en une série de trames identiques d'une longueur déterminée et une étape de filtrage fréquentiel (4) à l'aide d'un filtre de Wiener, caractérisé en ce qu'il comprend, en outre, les étapes suivantes :

- élaboration à partir desdits signaux bruités ($\{u(i)\}$) d'un modèle de bruit (1) sur un nombre N déterminé desdites trames, N étant compris entre des bornes minimale et maximale prédéterminées ;
- application d'une transformée de Fourier auxdites N trames ;
- estimation (2), pour chaque trame dudit modèle, de la densité spectrale de cette trame ;
- estimation (2) de la densité spectrale moyenne dudit modèle de bruit ;
- calcul (2), à partir de ces deux estimations, d'un coefficient de surestimation statistique, ledit coefficient statistique étant égal au rapport maximal, pour lesdites N trames du modèle de bruit, entre le maximum de la densité spectrale d'une trame considérée dudit modèle de bruit, et le maximum de la densité spectrale estimée du modèle de bruit ;
- estimation (3), pour chaque trame desdits signaux à débruiter ($\{u(i)\}$), de sa densité spectrale ; et ;
- modification (4), pour chaque trame desdits signaux à débruiter ($\{u(i)\}$), des coefficients dudit filtre de Wiener pour que la relation suivante soit vérifiée :

$$W(v) = \left(1 - \alpha \cdot \max_i \frac{\gamma_u(v)}{\gamma_w(v)} \right)^\beta,$$

relation dans laquelle α et β sont des coefficients fixes prédéterminés, dits coefficient statique de compensation énergétique et coefficient d'atténuation exponentielle, respectivement, v décrit l'ensemble des canaux fré-

quantiles de ladite transformée de Fourier, $\gamma_x(v)$ étant l'estimée de la densité spectrale de la trame à débiter, $\gamma_x(v)$ est ladite densité spectrale du modèle de bruit, et maxi ledit coefficient de surestimation statistique, modifiant le coefficient statique de compensation énergétique α .

2. Procédé selon la revendication 1, caractérisé en ce que ledit coefficient statistique *maxi* vérifie la relation suivante :

$$\text{maxi} = \text{Max}_{i=1..N} \left(\frac{\text{Max}(\gamma_i(v))}{\text{Max}(\gamma_x(v))} \right)$$

3. Procédé selon l'une des revendications 1 ou 2, caractérisé en ce qu'il comprend les étapes supplémentaires suivantes :

- calcul de l'énergie moyenne dudit modèle de bruit E_x ;
- calcul, pour chaque trame desdits signaux à débiter ($u(t)$), de l'énergie de la trame en cours E_y , et
- multiplication dudit coefficient statique de compensation énergétique α par un coefficient de pondération énergétique égal au rapport E_x/E_y , de manière à modifier sélectivement ces coefficients pour chaque trame desdits signaux à débiter ($u(t)$) par application d'un coefficient continuellement variable entre un extrême et un minimum, l'extrême étant sensiblement égal à l'unité lorsque lesdits signaux utiles sont absents desdits signaux à débiter ($u(t)$) et sensiblement égal à zéro lorsque l'énergie desdits signaux utiles est très supérieure à l'énergie desdits signaux de bruit, et à ce que lesdits coefficients du filtre de Wiener satisfassent la relation suivante :

$$W(v) = \left(1 - \alpha \cdot \text{maxi} \cdot \frac{E_x}{E_y} \cdot \frac{\gamma_x(v)}{\gamma_u(v)} \right)^\beta$$

4. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ledit coefficient statique de compensation énergétique α est égal à 10.

5. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ledit coefficient d'atténuation exponentielle β est égal à 0,5.

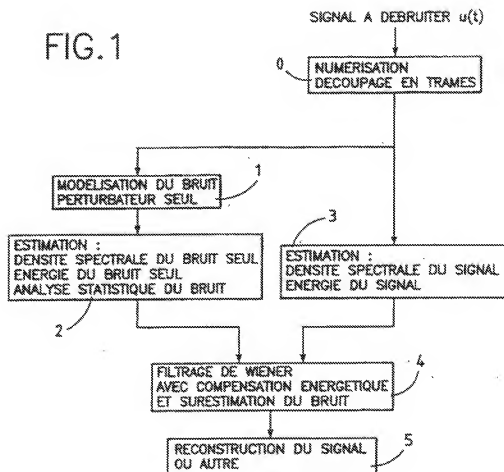
6. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend une étape initiale (0) consistant à numériser lesdits signaux à débiter ($u(t)$) par échantillonnage, chaque trame comprenant p échantillons.

7. Procédé selon la revendication 6, caractérisé en ce que ledit modèle de bruit (1) est obtenu par une recherche répétitive effectuée en permanence dans lesdits signaux à débiter ($u(t)$), en recherchant N trames successives, de p échantillons chacune, ayant les caractéristiques attendues d'un bruit, en stockant les $N \times p$ échantillons correspondants pour constituer ledit modèle de bruit, et en réitérant la recherche pour trouver un nouveau modèle de bruit et stocker le nouveau modèle en remplacement du précédent ou conserver le modèle précédent selon les caractéristiques respectives des deux modèles.

8. Application du procédé selon l'une quelconque des revendications précédentes au débruitage de signaux de parole brutes ($u(t)$).

9. Application du procédé selon la revendication 8, caractérisé en ce que la durée desdites trames est comprise dans la gamme 10 à 20 ms.

FIG.1



$$U(n) = S(n) + X(n) \rightarrow W(z) \rightarrow S(n)$$

FIG.2

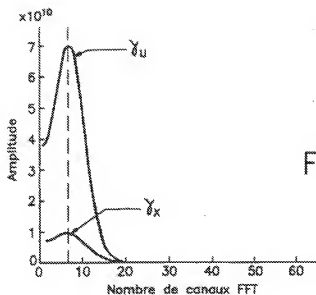


FIG. 3

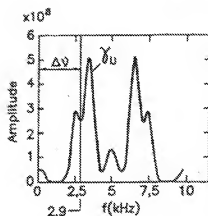


FIG. 4A

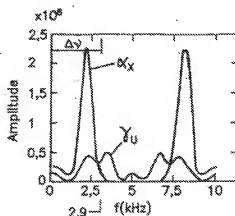


FIG. 4B

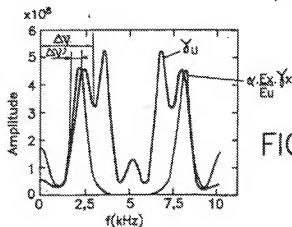


FIG. 5

FIG.6

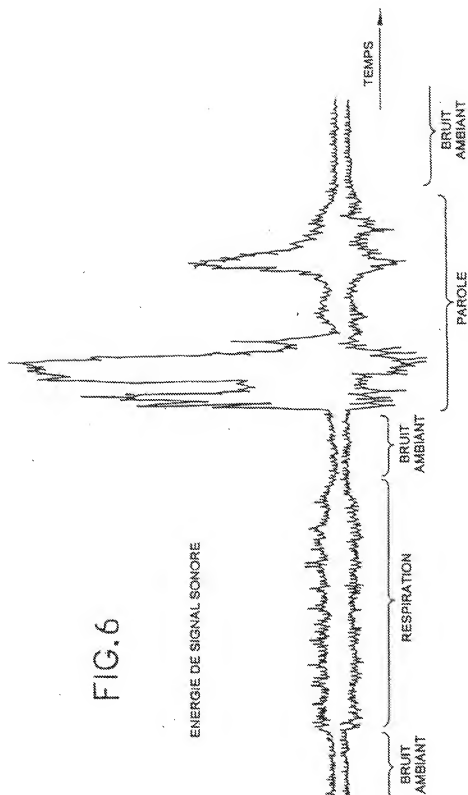


FIG.7

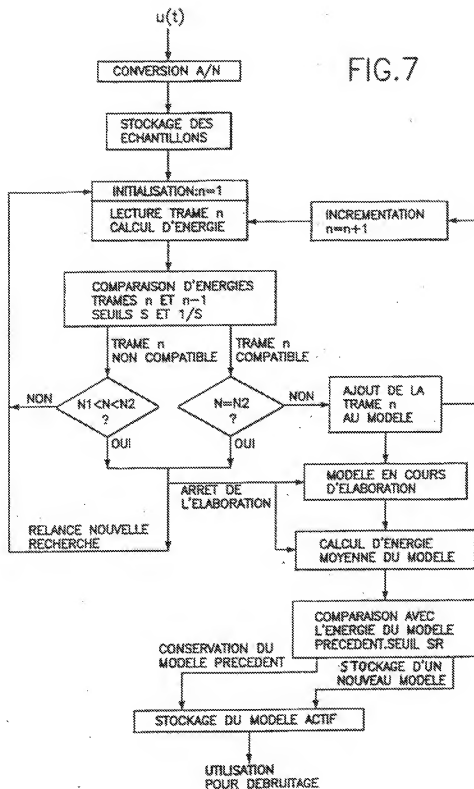
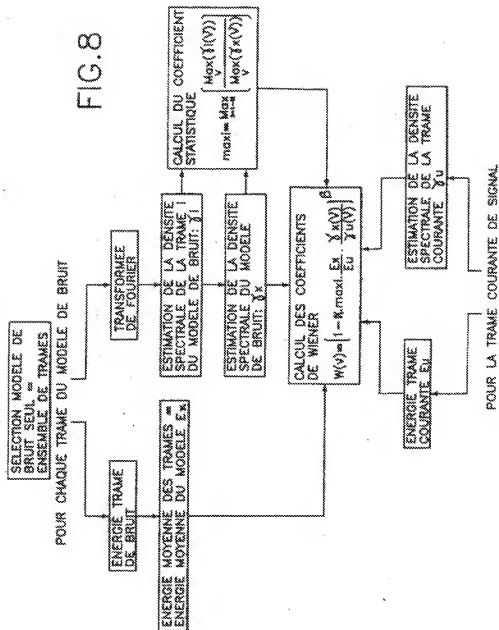


FIG.8





Office européen
des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numéro de la demande

EP 98 40 2894

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (M.C.L.B)
D,A	ARSLAN L ET AL: "NEW METHODS FOR ADAPTIVE NOISE SUPPRESSION" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), DETROIT, MAY 9 - 12, 1995 SPEECH, vol. VOL. 1, 9 mai 1995, pages 812-815, XP000658118 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS * alinéa 4 *	1-5,8,9	G10L3/02
A	HANSEN J H L ET AL: "Text-directed speech enhancement employing phone class parsing and feature map constrained vector quantization" SPEECH COMMUNICATION, vol. 21, no. 3, avril 1997, page 169-189 XP004059541 * page 180 *	1-5,8,9	
A	SUN T S ET AL: "SPEECH ENHANCEMENT USING A TERNARY-DECISION BASED FILTER" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), DETROIT, MAY 9 - 12, 1995 SPEECH, vol. VOL. 1, 9 mai 1995, pages 820-823, XP000658120 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS * figure 1 *	1-5,8,9	DOMAINES TECHNIQUES RECHERCHES (M.C.L.B) G10L
A	EP 0 534 837 A (MATRA COMMUNICATION SOCIETE AN) 31 mars 1993 * colonne 2, ligne 16 - ligne 20 * * figure 4 *	1-5,8,9	
Le présent rapport a été établi pour toutes les revendications			
Date de la recherche		Date d'établissement de la recherche	
LA HAYE		27 janvier 1999	
CATEGORIE DES DOCUMENTS CITES		Examinateur	
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : antérieur pour l'aspect technique O : divulgation non écrite P : document prioritaire		T : théorème ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons S : membre de la même famille, documents correspondants	

EP 0 918 317 A1 (1999-01-27)

**ANNEXE AU RAPPORT DE RECHERCHE EUROPEENNE
RELATIF A LA DEMANDE DE BREVET EUROPEEN NO.**

EP 98 40 2894

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche européenne visé ci-dessus.
Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

27-01-1999

Document brevet cité en rapport de recherche	Date de publication	Membres de la famille de brevets)	Date de publication
EP 0534837 A	31-03-1993	FR 2681715 A	26-03-1993
		FI 924287 A	26-03-1993

SPC.COM.FR/44

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No 1282